

You can submit the assignment in a pdf file, Jupyter notebook, or you can write down your discussion and explanation in Excel or python code, etc. You are welcome to use any software. Please make sure to submit the file that contains your analyses too (Excel file, R codes, etc.).

1 Commute versus non-commute hours

1. Do the commuting hours experience a higher number of ride-sharing trips? Is the difference statistically significant? What about the rate of using each ride-share service? (15%)
2. Assume on average riders pay around 12.5 dollars for a POOL ride and 10.5 on an Express ride. What is the difference in Uber's profit in a typical hour of commuting and non-commuting? Is the difference statistically significant? (15%)
3. Create a dashboard to plot histograms of ride-share trips from each service in one graph. Provide a simple graphical method for your user to filter commuting and non-commuting hours. (15%)
4. Using the kernel density estimation method with a suitable choice of bandwidth, estimate the distribution of cancellations rate in non-commuting and commuting hours. Do you observe a difference? How could you explain the difference? (15%)

2 Waiting times and commuting versus non-commuting hours

1. What is the difference in the number of ridesharing trips between the treatment and control groups during commuting hours? Is the difference statistically significant? What about rider cancellation and payouts? (20%)
2. What is the difference in the number of ridesharing trips between the treatment and control groups during non-commuting hours? Is the difference statistically significant? What about rider cancellation and payouts? (20%)